

WestminsterResearch

<http://www.westminster.ac.uk/westminsterresearch>

**Bridging the Gap Between Imaging Performance and Image
Quality Measures**

**Fry, E., Triantaphillidou, S., Jacobson, R., Jarvis, J. and Fagard-
Jenkin, R.**

This is a copy of a paper presented at IS&T Electronic Imaging Symposium 2018 -
Image Quality System Performance XV, San Francisco, CA, USA, 28 Jan - 1 Feb 2018.

It has been published in Electronic Imaging, DOI: 10.2352/ISSN.2470-
1173.2017.12.IQSP-255 and is available from the publisher at:

<http://ist.publisher.ingentaconnect.com/content/ist/ei/pre-prints/co...>

The WestminsterResearch online digital archive at the University of Westminster aims to make the
research output of the University available to a wider audience. Copyright and Moral Rights remain
with the authors and/or copyright owners.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely
distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

Bridging the Gap Between Imaging Performance and Image Quality Measures

Edward W. S. Fry¹, Sophie Triantaphillidou¹, Ralph E. Jacobson¹, John R. Jarvis¹, and Robin B. Jenkin²; ¹University of Westminster UK and ²NVIDIA USA

Abstract

Imaging system performance measures and Image Quality Metrics (IQM) are reviewed from a systems engineering perspective, focusing on spatial quality of still image capture systems. We classify IQMs broadly as: Computational IQMs (CP-IQM), Multivariate Formalism IQMs (MF-IQM), Image Fidelity Metrics (IF-IQM), and Signal Transfer Visual IQMs (STV-IQM). Comparison of each genre finds STV-IQMs well suited for capture system quality evaluation: they incorporate performance measures relevant to optical systems design, such as Modulation Transfer Function (MTF) and Noise-Power Spectrum (NPS); their bottom-up, modular approach enables system components to be optimized separately. We suggest that correlation between STV-IQMs and observer quality scores is limited by three factors: current MTF and NPS measures do not characterize scene-dependent performance introduced by imaging system non-linearities; contrast sensitivity models employed do not account for contextual masking effects; cognitive factors are not considered. We hypothesize that implementation of scene and process-dependent MTF (SPD-MTF) and NPS (SPD-NPS) measures should mitigate errors originating from scene dependent system performance. Further, we propose implementation of contextual contrast detection and discrimination models to better represent low-level visual performance in image quality analysis. Finally, we discuss image quality optimization functions that may potentially close the gap between contrast detection/discrimination and quality.

Introduction

This paper is concerned with spatial image capture system performance measurement, spatial image quality modeling, and their relationship. Both involve spatial luminance contrast signals, which are principle carriers of information that is fundamental to human understanding of shape form and detail, and which contribute heavily toward the overall perceived image quality.

Perceived image quality can be evaluated directly by psychophysical tests involving human observers. These tests are time consuming, expensive, and prone to inaccuracy with incorrect implementation. In parallel, performance measures such as the *Modulation Transfer Function* (MTF) and *Noise-Power Spectrum* (NPS) are used to characterize imaging system behavior. There are well-established assumptions about functional relationship(s) between these measures and overall subjective system image quality, or the quality of individual image attributes. Finally, image quality metrics (IQM) produce output scores that aim to correlate with perceived quality or fidelity.

Linear system theory is core to MTF, NPS and any IQMs employing them as input parameters. It requires systems analyzed to be linear, spatially invariant and homogenous [1]. Today's capture systems often apply non-linear content-aware processing algorithms. These introduce artefacts and alter structural content depending upon original scene information. This poses significant challenges to capture system characterization and quality

modelling, as the targets and techniques used affect the results obtained. This paper reviews spatial performance measures and relevant IQMs from a capture systems engineering perspective, with this issue in mind. It further classifies IQMs into four genres, which we denote as: *Computational IQMs* (CP-IQM), *Multivariate Formalism IQMs* (MF-IQM), *Image Fidelity Metrics* (IF-IQM), and *Signal Transfer Visual IQMs* (STV-IQM). Frameworks are finally proposed to revise system performance measures and selected relevant IQMs, rendering them more appropriate for non-linear content-aware capture systems.

Spatial System Performance Review

The MTF and NPS are standard measures of spatial signal transfer and noise performance, respectively. They are both based on Fourier theory of image formation [1].

Noise-Power Spectrum (NPS)

The NPS characterizes a system's additive noise, and thus informs upon system noisiness. Equation 1 defines NPS, where $g(x, y)$ is the intensity of the captured image at x, y coordinates, and $\bar{g}(x, y)$ is its expectation values (i.e. the mean signal). Noise is commonly measured from uniform luminance patches, rendering $\bar{g}(x, y)$ constant across all pixel locations. Uniform patches characterize linear system noise accurately, but are not necessarily suitable when characterizing non-linear content-aware systems.

$$NPS(u, v) = |DFT(g(x, y) - \bar{g}(x, y))|^2 \quad (1)$$

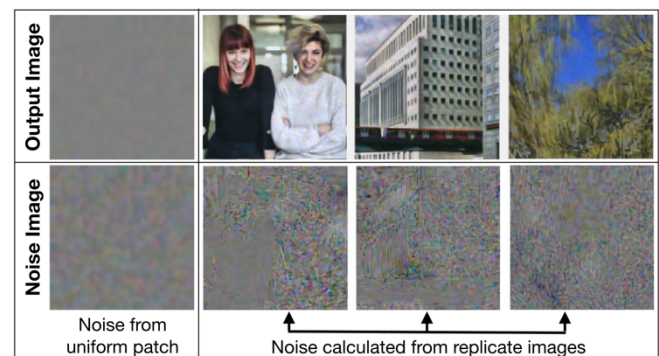


Figure 1. Noise images generated from a camera simulation pipeline. Poisson noise was generated at a linear signal-to-noise ratio of 10. Non-linear content-aware denoising and sharpening were then applied. The noise image obtained by processing a uniform patch (left) underestimates the noise present in pictorial images (right). Noise is measured from the latter by subtracting the mean of ten replicate pictorial images, from a single frame, as per Figure 6. Noise image contrast is enhanced to demonstrate noise scene-dependency.

Non-linear content-aware denoising filters are a common source of image capture system non-linearity. They preserve image structure and edge content by making local adjustments to their processing intensity, according to signal gradient characteristics. Since local image structure impedes their denoising ability, the intensity and

spatial distribution of an image's noise, after filtration, is scene-dependent (Figure 1). Uniform patches provide the easiest scenario for such algorithms to operate in, and resultant NPS measurements underestimate real-world noise power. Later in this paper, we propose a *scene and process-dependent NPS* (SPD-NPS) measure that addresses this issue, by measuring noise from pictorial images.

Modulation Transfer Function (MTF)

The MTF (and the related *Spatial Frequency Response* (SFR) [2]) characterize a system's reproduction of modulation, as a function of spatial frequency. MTF informs upon system resolution and sharpness, which are major contributors to overall quality. It is traditionally measured from sinusoid, edge and noise targets. For non-linear systems, however, no unique MTF exists, and the above methods deliver different results. Sinusoid and edge targets describe limiting resolution and subjective sharpness respectively for these systems [3], but do not describe perceived texture loss effectively [4]. Texture loss is a primary driver of overall perceived image quality in today's non-linear capture systems [5].

Dead leaves MTFs provide useful measurements of texture loss, and general spatial performance of non-linear content-aware systems. They do so by replicating statistical properties of natural scenes to trigger non-linear content-aware processes at 'normal' levels [4]. The direct dead leaves implementation [6] (Equation 2) improves accuracy by subtracting the system's NPS measured from a uniform patch ($NPS_{Output}(u)$), from the output target power spectrum ($PS_{Output}(u)$); $PS_{Input}(u)$ is the input target power spectrum; u is the spatial frequency. A further intrinsic implementation characterizes performance of the lens and imager; it is not sensitive to reversible image processing such as sharpening or contrast stretching [7], which affect perceived image quality.

$$MTF(u) = \sqrt{\frac{PS_{Output}(u) - NPS_{Output}(u)}{PS_{Input}(u)}} \quad (2)$$

The dead leaves target does not characterize non-linear content-aware system performance with respect to a given scene. The *scene and process-dependent MTF* (SPD-MTF) measure proposed by Branca et al. [8] achieves this by substituting pictorial image power spectra into the direct dead leaves calculation, to trigger non-linear content-aware processing appropriately for each scene. Initial results suggest that real capture systems display scene-dependent signal transfer performance. This was shown to be greater in measurements from a smartphone than a professional DSLR camera, due to the former's increased non-linear content-aware processing [8]. As with the direct dead leaves MTF, this version of the SPD-MTF calculation accounts for NPS measured from uniform patches, which introduces error. The SPD-MTF framework proposed later in this paper mitigates this problem by implementing a revised and more suitably measured SPD-NPS.

A Review of Image Quality Metrics (IQM)

IQMs are used to evaluate image or imaging system quality, without need for psychophysics (Figure 2). They map objective input data (images or system performance measurements) to output scores that aim to correlate with impressions of quality, quality difference or fidelity (perceived magnitude of image differences).



Figure 2. Flow-chart describing image quality metrics (IQM) in general terms.

Image quality modelling forms a converging point of several research areas: neural physiology, visual psychophysics, computer vision, machine learning, signal and image processing and imaging systems engineering. Specialists have approached IQM development from different angles, and a broad spectrum of IQMs has evolved, with various advantages and applications.

Classification of Image Quality Metrics (IQM)

IQM classifications are summarized below. Univariate and artefact-specific IQMs predict quality with respect to a single image quality attribute or artefact, respectively. Multivariate IQMs measure quality with respect to two or more quality attributes.

Visual modelling differs greatly between IQMs. Bottom-up approaches model separate low-level visual processes in a step-by-step fashion with direct reference to underlying physics, and psychophysical data. Top-down approaches model the combined effect of low-level vision and high-level cognitive processing in one step, or as a combination of black box sub-processes.

IQMs that require images as input data are categorized as no-reference, full-reference and reduced-reference. No-reference IQMs require data from a single output image, and correlate with *Mean Opinion Scores* (MOS) or quality scores in just-preferable difference (JPD) units. Full-reference and reduced-reference IQMs compare data from a test and reference image, where the former is a processed version of the latter. Full-reference IQMs require all information from both images. Reduced-reference IQMs require certain features only. Full-reference IF-IQMs measure fidelity and either correlate with probabilities for image difference detection or fidelity scores in units of just-noticeable difference (JND). Full-reference and reduced-reference IQMs designed to predict quality correlate with *Differential Mean Opinion Scores* (DMOS) or quality scores related by JPD. This paper classifies IQMs into four further genres, which we name: STV-IQMs, CP-IQMs, MF-IQMs and IF-IQMs. Each genre is summarized in Table 1, and is reviewed later from a capture systems engineering perspective.

Table 1: Summary of IQM genres

Metric Genre	Reference Type	Input Parameter	Output Data	Correlating Data
STV-IQM	No Reference	Test chart measurements	Quality or sharpness score	Perceived quality or sharpness score (JPD)
MF-IQM	No Reference	Univariate metric scores, from test chart measurements	Quality score	Perceived quality score (JPD)
IF-IQM	Full Reference	Test image + Reference image	Fidelity score or error map	Perceived fidelity score (JND), or probability of difference visibility
CP-IQM	Full / Reduced Reference	Test image + Reference image	Quality difference score	Differential Mean Opinion Score (DMOS)
	No Reference	Test image	Quality score	Mean Opinion Score (MOS)

Signal Transfer Visual Image Quality Metrics (STV-IQM)

STV-IQMs were traditionally useful when engineering analog and early digital systems, and have been reviewed extensively [9]. They combine imaging system performance measures and bottom-up visual modelling, as per linear system theory. Input parameters may include MTF, NPS, scene DFT power spectra, and models of vision incorporating viewing conditions.

Univariate STV-IQMs such as the *Square Root Integral* (SQRI) [10] and the *Modulation Transfer Acutance* (AMTA) [11]

model perceived sharpness according to the system's signal transfer characteristics, and the viewing conditions. They commonly use an integrated cascade calculation, as generalized in Equation 3, where $MTF_{system}(u)$ is the system's MTF, $CSF(u)$ is a visual contrast sensitivity model, u is the retinal spatial frequency and u_{min} and u_{max} are the minimum and maximum perceptible frequencies, respectively.

$$Sharpness Score = \int_{u_{min}}^{u_{max}} [MTF_{system}(u) \cdot CSF(u)] du \quad (3)$$

Multivariate STV-IQMs such as *SQRI with Noise* (SQRIin) [12], *Pictorial Information Capacity* (PIC) [13] and *Effective Pictorial Information Capacity* (EPIC) [14] consider additive noise as well as system signal transfer characteristics. Most model image quality as the integral of a visually weighted signal-to-noise ratio (SNR), where the numerator weights scene content and system signal transfer parameters with a CSF, and the denominator consists of a visual noise model and visually weighted system noise parameter. Equation 4 generalizes these common multivariate STV-IQM calculations, where $PS_{scene}(u)$ is the scene's DFT power spectrum, $MTF_{system}(u)$ is the system's MTF, $NPS_{system}(u)$ is the system's NPS, $NPS_{visual}(u)$ is the neural noise, $CSF(u)$ is the contrast sensitivity function, u is the retinal spatial frequency and u_{min} and u_{max} are the minimum and maximum perceptible frequencies, respectively.

$$Quality Score = \int_{u_{min}}^{u_{max}} \left[\frac{PS_{scene}(u) \cdot MTF_{system}^2(u) \cdot CSF^2(u)}{NPS_{system}(u) \cdot CSF^2(u) + NPS_{visual}(u)} \right]^{0.5} \frac{du}{u} \quad (4)$$

Multivariate Formalism (MF-IQM)

MF-IQMs model image quality as a Minkowski combination of the perceived strength of individual quality attributes [15], where the most significant attribute dominates the overall score. Each attribute is modelled by a univariate metric in units of JND.

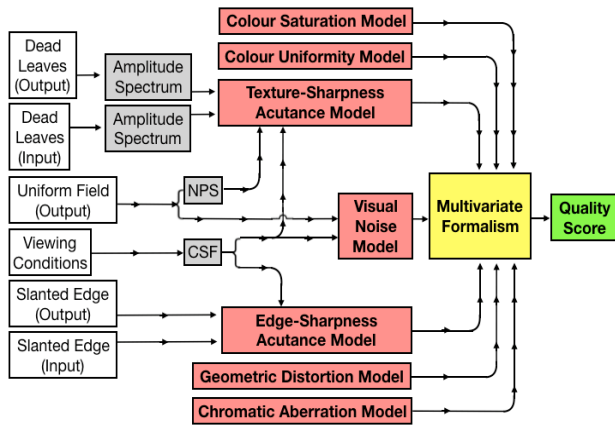


Figure 3. Flow chart summarizing the CPIQ group's MF-IQM [5]. Input parameters are specified for spatial signal transfer and noise components.

The Camera Phone Image Quality (CPIQ) group's MF-IQM [5] models the combined effect of color saturation, color uniformity, geometric lens distortion, chromatic aberration, edge-sharpness, texture-sharpness, and visual noise. Texture-sharpness and edge-sharpness metrics employed are univariate acutance STV-IQMs, implementing direct dead leaves [6] and ISO 12233 slanted-edge [2] MTFs respectively. The visual noise metric follows Annex B of ISO 15739 [16] and uses uniform patch noise measurements.

Figure 3 summarizes the CPIQ's MF-IQM, including input parameters for the above spatial signal transfer and noise metrics.

Image Fidelity Metrics (IF-IQM)

IF-IQMs are inherently full-reference, and model the perceptibility of differences between distorted (test) and optimum (reference) images. Threshold IF-IQMs model the probability for discrimination of such differences. Suprathreshold IF-IQMs model their perceived magnitude in units of JND.

Threshold IF-IQMs [17] such as the *Visible Difference Predictor* (VDP) [18] were developed for compression, imaging systems and display optimization applications. Initial processing often accounts for display characteristics, amplitude non-linearity, divisive normalization, visual masking and contrast sensitivity. Differences between test and reference images are then calculated, and probability models are applied. Output distortion maps state the probability for threshold discrimination of image differences at each pixel, and can be pooled to single figures using different techniques.

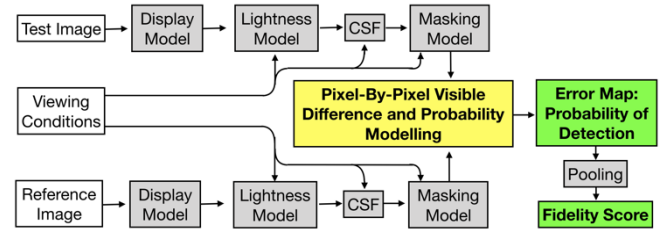


Figure 4. Generalization of threshold IF-IQM processing.

Suprathreshold IF-IQMs such as *S-CIELAB* [19] and the *modular image difference model* [20] filter opponent color space images with CSFs or suprathreshold contrast discrimination models. JND calculations are generally performed using standard color difference/appearance models, which may account for local contrast detection and chromatic adaptation [20]. The resultant image difference maps can also be pooled to single figures.

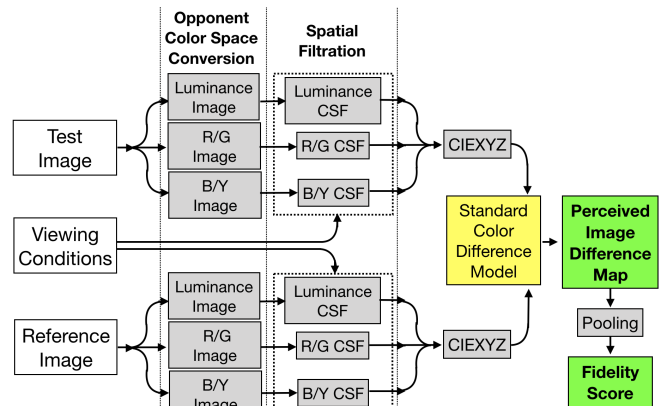


Figure 5. Generalization of suprathreshold IF-IQM processing.

Computational Metrics (CP-IQM)

CP-IQMs are the most commonly published IQMs of the past decade and have been reviewed frequently, e.g. [21]. They establish relationships between MOS/DMOS scores and image features or natural scene statistics (NSS), and usually implement top-down visual modelling. Full-reference and reduced-reference examples correlate with DMOS, and include mathematical, structural, visual attention, information theoretical, hybrid and

machine learning approaches. No-reference examples correlate with MOS, and include information theoretical and machine learning approaches. Each is summarized below.

Early mathematical approaches such as *Mean Square Error* (MSE) are most popular due to their simplicity and mathematical tractability. But they lack human visual system (HVS) modelling, and are less perceptually relevant than more complex CP-IQMs.

Since the introduction of the *Structural Similarity* (SSIM) index [22], structural IQMs have become the most widely accepted full-reference CP-IQMs. They assume the HVS is highly sensitive to structural distortions, defined as loss of correlation between test and reference images, after accounting for differences in their luminance and variance (contrast). Later revisions of SSIM implement multi-scale [23] and wavelet domain [24] processing.

Visual attention approaches [25, 26] weight image content with eye-tracking data, prior to the implementation of CP-IQMs such as MSE or SSIM. This generally improves accuracy, but can cause issues when relevant distortions are in non-salient areas [26]. Implementation of algorithmically generated saliency maps instead increases computational complexity and is less accurate [26].

Information theoretical approaches such as *Visual Information Fidelity* (VIF) [27] and the *Information Fidelity Criterion* (IFC) [28] share foundations in information theory with some STV-IQMs [13, 14]. They differ from the latter since their visual modelling is top-down. This accounts for divisive normalization and orientational selectivity, but not the CSF or luminance masking.

Hybrid approaches generally combine bottom-up low-level visual models with black box high-level vision models [29, 30]. Others extend existing CP-IQMs to account for low-level vision.

Machine learning approaches [31, 32] consider image quality assessment as a supervised regression problem between images and MOS/DMOS scores. Features are extracted to narrow-down the input data. These are commonly distortion specific, such as ringing, blocking, texture, blurring and noise. Several recent examples solve this regression task using Convolutional Neural Networks (CNN), and include unsupervised approaches [33].

A Capture Systems' Engineering Perspective of IQMs

MF-IQMs are used widely in capture system image quality assessment, and feature in recent ISO 20462 image quality ruler [34] and CPIQ IQM [5] standards. They are modular, adaptable and provide logical pathways towards their quality scores. They employ standard system performance measures and mechanistic bottom-up visual models that account for viewing conditions, and thus can control system performance characteristics, relative to the perceived image quality. The CPIQ's implementation (Figure 3) is more complex to execute and computationally intensive than STV-IQMs, or CP-IQMs. Out of its seven univariate IQM components, those relating to spatial signal transfer and visual noise are most significant to overall perceived quality [5]. Both these components are covered by multivariate STV-IQMs. When texture-sharpness models are employed, issues regarding scene-dependent system performance are possibly mitigated, but edge-sharpness and visual noise models suffer from them. The employed CSFs do not account for visual masking present in image viewing.

IF-IQMs are modular cascade-based calculations that also provide logical pathways towards their fidelity scores, and implement mechanistic bottom-up HVS models that account for viewing conditions. Comparing test and reference images on a pixel-to-pixel level allows them to implement more extensive 2D HVS models (e.g. 2D CSFs), and increases computational workload. Reference images, however, do not exist in capture system image quality assessment, and standard system

performance measures are unaccounted for. Perceived fidelity and quality also correlate only to some extent. No universal method exists for pooling distortion maps into single figures. Threshold IF-IQMs generally clip suprathreshold image differences rendering them unsuitable for some image quality assessment applications.

CP-IQMs fulfill the increasing demand for accurate IQMs that are simple to implement (including in real-time) with minimal input parameters. They are rarely used in capture systems image quality assessment, for the following reasons. Standard system performance measures are not included, and thus a link between system design and perceived quality is not readily available. Their top-down visual models are generally non-mechanistic, where low-level and high-level components consist of black box sub-components that often have limited psychophysical basis, or rely on domain specific knowledge. Also, they generally do not account for the viewing conditions. No-reference CP-IQMs are better suited for capture system quality assessment than full and reduced-reference CP-IQMs, although they are generally less modular, and often have minimum, or no relation to the underlying system's physics.

STV-IQM characteristics are most suitable for capture systems' image quality assessment, alongside MF-IQMs. Their visual modelling makes them more relevant to image quality assessment than assumptions based upon MTF or NPS measures alone, whilst they still relate directly to each performance measure. As with MF-IQMs and IF-IQMs, quality scores are computed in logical and modular step-by-step calculations, where input parameters are (or relate directly to) real physical quantities. Thus, they model the relationship between perceived image quality and its underlying physics, establishing causal justification for their quality scores. Input parameters can be altered independently, enabling imaging chain variations to be simulated, under different viewing conditions. As with MF-IQMs, STV-IQMs are reliant, however, upon traditional 1D system performance measures and visual models which do not fully characterize the relevant imaging system or visual perceptive processes. STV-IQMs have been most successful when modelling linear system image quality. They encounter problems when assessing non-linear processes, which can create visually unpleasant artefacts that are unaccounted for, or contribute positively to metric scores. Although STV-IQMs may correlate well with subjective ratings for individual images from such non-linear systems, overall correlation is generally poor [9].

Based on the above, we hypothesize that some inaccuracies may result from the following contributing factors:

1. MTF and NPS measures currently employed do not account for scene-dependency in imaging system performance.
2. CSFs currently employed do not accurately describe visual performance relevant to image quality assessment tasks.
3. Cognitive factors are not taken into account.

Factor 1 has been expanded previously, during the system performance measure review. Factors 2 and 3 are discussed below.

STV-IQMs apply 1D CSFs directly as weighting functions, thus assuming they are HVS transfer functions. Decades of successful STV-IQM usage suggests this assumption may be valid to some extent, although HVS linearity only applies near to threshold limits [35]. The question whether image quality is a function of threshold or suprathreshold vision has been debated; it is likely that both processes are involved [36]. In either case, the analyzed image signals are complex, and viewed in the context of other image content, thus introducing contrast and texture masking effects. We propose that *Contextual Contrast Sensitivity Functions*

(cCSFs - threshold models), and *Contextual Visual Perception Functions* (cVPFs - suprathreshold models) [36] are more appropriate models than the CSFs currently employed, which are based upon viewing isolated simple stimuli. cCSF and cVPF describe detection and discrimination of real scene content respectively, and account for the visual masking that results from viewing complex image signals [36].

Cognitive processing – which accounts for observer quality consciousness and subjective preference – is unaccounted for by CSFs. Cognitive processing tends to prioritize the importance of certain suprathreshold signals above others during subjective IQA tasks. Signals of importance may be from the image itself, or its artifacts. These effects are usually scene-dependent.

In the following section, we propose a revised STV-IQM framework that accounts for Factors 1 and 2; A method to account for Factor 3 is briefly discussed in the conclusions.

Proposed Framework

We propose a STV-IQM framework, in which the input parameters (i.e. NPS, MTF, CSF) represent the response of the imaging and visual systems to real, complex images, rather than simple test charts or stimuli.

Scene and Process-Dependent NPS (SPD-NPS)

The proposed SPD-NPS accounts for non-linear, content-aware processes at the expense of increased computational processing. SPD-NPS is ideally measured from pictorial images. Alternatively, it can be measured from the dead leaves target to approximate the system's noise performance for an 'average scene spectrum'. Our initial simulations using non-linear content-aware camera pipelines indicate that both SPD-NPS variants are more representative image noise measures than NPSs measured from uniform patches. SPD-NPS is described in Figure 6, and calculated as per Equation 1, where $g(x, y)$ is the output image, and $\bar{g}(x, y)$ is the mean output of several individual output image replicates. A rotational average of 2D SPD-NPS yields the 1D SPD-NPS. All images must be accurately registered. *Fixed Pattern Noise* (FPN) is unaccounted for, but is less significant in contemporary capture systems under most conditions, and can be measured separately.

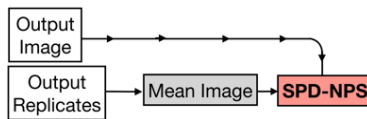


Figure 6. Flow-chart describing the SPD-NPS framework.

Scene and Process-Dependent MTF (SPD-MTF)

The proposed SPD-MTF expands upon the original SPD-MTF measurement presented in [8], where the system MTF is measured directly from natural scenes. SPD-NPS (Figure 6) is implemented to account for scene-dependent processing of noise during the signal transfer calculation. Our initial simulations using non-linear content-aware camera pipelines indicate SPD-MTFs are more representative of system performance than direct dead leaves MTFs.

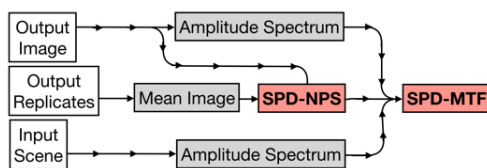


Figure 7. Flow-chart describing the revised SPD-MTF framework.

SPD-MTF is described in Figure 7 and calculated as per Equation 2, where $NPS_{Output}(u)$ is SPD-NPS and $PS_{Input}(u)$ and $PS_{Output}(u)$ are input and output pictorial image power spectra.

Revised Multivariate STV-IQM Framework

Existing multivariate STV-IQM frameworks (as generalized in Equation 4) are revised as per the following changes. SPD-MTF and SPD-NPS replace the imaging system MTF and NPS derived from test charts. Both are expected to reduce STV-IQM inaccuracy due to imaging system scene-dependency (Factor 1, above). cCSF or cVPF [36] functions (or a combination of both) are substituted in place of the traditional CSF. Their inclusion should reduce STV-IQM inaccuracy due to lower-level HVS model limitations (Factor 2, above). Figure 8 describes this revised STV-IQM framework.

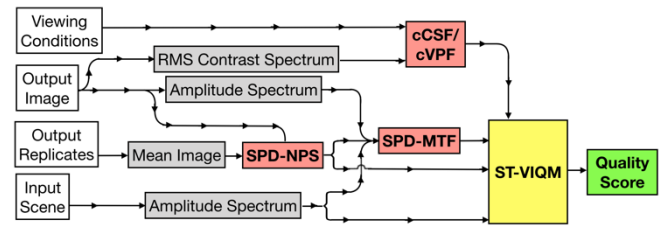


Figure 8. Flow-chart describing the revised multivariate STV-IQM framework.

Conclusions

Imaging system performance measures and genres of IQMs are reviewed from a capture systems' engineering perspective. The latter review concludes that multivariate STV-IQMs have advantageous characteristics, but require updating. SPD-NPS and SPD-MTF measures, which account for non-linear content-aware processing are thus proposed to replace the traditionally measured NPS and MTF input parameters. These measures form the basis of the revised STV-IQM framework, alongside cCSF and cVPF visual models that account for effects of masking on contrast detection and discrimination. The proposed measuring techniques and IQM framework are currently in the process of validation.

Finally, in order to fully describe image quality perception, STV-IQMs will need to account for high-level cognitive processing, associated with observer quality consciousness and subjective preference characteristics. This is a real challenge since these factors are scene and observer-dependent. We propose the incorporation of *Optimal Contrast Weighting* (OCW) functions [37] in STV-IQMs to describe the required frequency domain individual image/observer optimization of the visual quality. Optimization is achieved by cascading high-pass filters with scene-dependent cCSF/cVPF functions, thus only boosting certain detected and discriminated signals. Optimized images contain slightly higher than natural levels of mid-to-high frequency contrast, thus enhancing perceived sharpness and structure. OCWs do not account for the presence of suprathreshold image artefacts, or their impact on the quality evaluation process. They are currently being investigated further, using a purpose-built image band contrast equalizer [38], and can be implemented as parameters in the STV-IQMs presented in this paper, to model relevant cognitive processes concerning spatial image quality judgement. This is with the intention of further closing the gap between system performance measurement, contrast detection/discrimination modelling and image quality perception.

Abbreviations

SPD-MTF..... Scene and Process-Dependent MTF
 SPD-NPS..... Scene and Process-Dependent NPS

STV-IQM..... Signal Transfer Visual Image Quality Metric
 MF-IQM..... Multivariate Formalism Image Quality Metric
 IF-IQM..... Image Fidelity Metric
 CP-IQM..... Computational Image Quality Metric
 cCSF..... Contextual Contrast Sensitivity Function
 cVPF..... Contextual Visual Perception Function
 OCW..... Optimal Contrast Weighting Function

Author Biography

Edward Fry is a PhD. student at the University of Westminster, UK, in the field of imaging systems performance measurement and image quality modeling. His research focuses on scene-dependency in imaging systems performance, and its impact upon overall perceived image quality.

References

- [1] J. C. Dainty and R. Shaw, *Image Science: Principles, Analysis and Evaluation of Photographic-type Imaging Processes*. London: Academic Press Ltd., 1974.
- [2] ISO 12233:2014, *Photography - Electronic still picture imaging - Resolution and spatial frequency responses*. Geneva: International Organization for Standardization, 2014.
- [3] L. Kirk, P. Herzer, U. Artmann, and D. Kunz, "Description of Texture Loss Using the Dead Leaves Target - Current Issues and a New Intrinsic Approach," in *Proc. SPIE 9023, Digit. Photog. X*, 2014.
- [4] F. Cao, F. Guichard, and H. Hornung, "Measuring Texture Sharpness of A Digital Camera," *Proc. SPIE 7250, Digit. Photogr. V*, 2009.
- [5] E. W. Jin et al., "Towards the Development of the IEEE P1858 CPIQ Standard - A validation study," *Proc IS&T Image Qual. Syst. Perform. XIV*, pp. 88–94, 2017.
- [6] J. McElvain, S. P. Campbell, J. Miller, and E. W. Jin, "Texture-based measurement of spatial frequency response using the dead leaves target: extensions, and application to real camera systems," in *Proc. SPIE 7537, Digit. Photog. VI*, 2010, p. 75370D–1.
- [7] P. D. Burns and J. Martinez Bauza, "Intrinsic camera resolution measurement," *Proc. SPIE 9396, Image Qual. Syst. Perform. XII*, vol. 9396, p. 939609, 2015.
- [8] R. Branca, S. Triantaphillidou, and P. Burns, "Texture MTF from images of natural scenes," *Proc. IS&T Image Qual. Syst. Perform. XIV*, pp. 113–120, 2017.
- [9] R. Jacobson and S. Triantaphillidou, "Metric Approaches to Image Quality," in *Colour image science: exploiting digital media*, John Wiley & Sons, 2002, pp. 371–392.
- [10] P. G. J. Barten, "Evaluation of subjective image quality with the square-root integral method," *J. Opt. Soc. Am. A*, vol. 7, no. 10, p. 2024, 1990.
- [11] E. M. Crane, "Acutance and Granulance," *Proc. SPIE Image Qual. Vol. 310*, pp. 125–132, 1981.
- [12] P. G. J. Barten, "Evaluation of the Effect of Noise on Subjective Image Quality," in *Proc. SPIE 1453, Human Vision, Visual Processing, and Digital Display II*, 1991, vol. 1453, pp. 2–15.
- [13] K. Topfer and R. E. Jacobson, "The Relationship Between Objective and Subjective Image Quality Criteria," *J. Inf. Rec. Mats.*, vol. 21, pp. 5–27, 1993.
- [14] R. B. Jenkin, S. Triantaphillidou, and M. A. Richardson, "Effective Pictorial Information Capacity as an Image Quality Metric," *Proc. SPIE 6494, Image Qual. Syst. Perform. IV*, 2007.
- [15] B. W. Keelan, *Handbook of Image Quality, Characterisation and Prediction*. New York: Marcel Dekker, 2002.
- [16] ISO 15739:2013, *Photography — Electronic still-picture imaging — Noise measurements*. 2013.
- [17] J. Lubin, "The Use of Psychophysical Data and Models in the Analysis of Display System Performance," in *Digital Images and Human Vision*, A. Watson, Ed. MIT Press, 1993, pp. 163–178.
- [18] S. Daly, "The visible difference predictor: an algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, A. B. Watson, Ed. The MIT Press, 1993, pp. 179–206.
- [19] X. Zhang and B. A. Wandell, "A spatial extension of CIELAB for digital color image reproduction," *J. Soc. Inf. Disp.*, vol. 5, no. 1, p. 61, 1997.
- [20] M. D. Fairchild and G. M. Johnson, "iCAM framework for image appearance, differences, and quality," *J. Electron. Imaging*, vol. 13, no. 1, p. 126, 2004.
- [21] M. Pedersen and J. Y. Hardeberg, "Survey of full-reference image quality metrics," *Høgskolen i Gjøviks Rapp. 5, Nor. Color Res. Lab. Gjøvik Univ. Coll.*, no. 5, pp. 1–74, 2009.
- [22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.
- [23] Z. Wang and A. C. Bovik, "Multi-Scale Structural Similarity For Image Quality Assessment," *Proc. IEEE Asilomar Conf. Signals*, vol. 2, no. Ki L, pp. 1398–1402, 2003.
- [24] Z. Wang and E. P. Simoncelli, "Translation Insensitive Image Similarity in Complex Wavelet Domain," *IEEE Int. Conf. Acoust. Speech, Signal Process.*, pp. 573–576, 2005.
- [25] X. Min, G. Zhai, Z. Gao, and K. Gu, "Visual attention data for image quality assessment databases," *2014 IEEE Int. Symp. Circuits Syst.*, pp. 894–897, 2014.
- [26] H. Liu and I. Heynderickx, "Visual attention in objective image quality assessment: Based on eye-tracking data," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 7, pp. 971–982, 2011.
- [27] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, 2006.
- [28] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, 2005.
- [29] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *J. Electron. Imaging*, vol. 19, no. 1, p. 11006, 2010.
- [30] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, 2007.
- [31] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: from natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–64, 2011.
- [32] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, 2012.
- [33] P. Ye, J. Kumar, D. Doermann, and L. Kang, "Unsupervised feature learning framework for no-reference image quality assessment," *IEEE Comput. Vis. Pattern Recognit.*, pp. 1098–1105, 2012.
- [34] ISO 20462-3:2012, *Photography -- Psychophysical experimental methods for estimating image quality -- Part 3: Quality ruler method*.
- [35] P. G. J. Barten, *Contrast sensitivity of the human eye and its effects on image quality*. SPIE press, 1999.
- [36] S. Triantaphillidou, J. Jarvis, and G. Gupta, "Spatial contrast sensitivity and discrimination in pictorial images," *Proc. SPIE Vol. 9016 Image Qual. Syst. Perform. XI*, 2014.
- [37] E. W. S. Fry, S. Triantaphillidou, J. Jarvis, and G. Gupta, "Image quality optimization, via application of contextual contrast sensitivity and discrimination functions," *Proc. SPIE 9396, Image Qual. Syst. Perform. XII*, 2015.
- [38] S. Triantaphillidou, E. W. S. Fry, V. Sanchis-Jurado, and A. Pons, "Image Quality Loss and Compensation for Visually Impaired Observers," in *Proc IS&T Image Qual. Syst. Perform. XV*, 2017.